

УДК 004.724.4(045)

Давиденко І.М., канд. техн. наук
Воевода Д.С.

АЛГОРИТМ ФОРМУВАННЯ СТІЙКОЇ СТРУКТУРИ МОБІЛЬНОЇ GRID-СИСТЕМИ

Інститут комп'ютерних технологій
Національного авіаційного університету

Запропоновано алгоритм формування стійкої структури мобільної GRID-системи з забезпеченням заданих параметрів якості обслуговування та з урахуванням нестабільності підключення мобільних вузлів, що дало можливість підвищити продуктивність системи

Вступ

У зв'язку з динамічним розвитком мобільного середовища актуальним є питання визначення кола послуг, які необхідно забезпечувати у такій структурі. Крім того актуальними стають питання нестабільного підключення, неоднорідності пристрою та слабка безпека. Серед цих питань найважливішими є проблеми перебою під час роботи ресурсу, оскільки мобільні ресурси схильні до частих відключень через обмежений діапазон зв'язку і часту зміну місце розташування. У зв'язку з цим актуальним є розробка алгоритму формування стійкої структури мобільної GRID-системи.

Огляд і аналіз існуючих рішень

У динамічній GRID-системі керування ресурсами відбувається розподілено та автономно [1], утворюючи обчислювану мережу, яка може переносити та виконувати величезну кількість роботи з точки зору швидкості обчислень і продуктивності. В роботі [2] ретельно розглянуто способи та алгоритми інтеграції мобільних та мережевих систем, у результаті чого зроблено висновок щодо функції мобільних пристроїв в GRID-системах:

1. Мобільні пристрої можуть бути використані в якості інтерфейсу GRID, що надасть можливість ініціювати використання ресурсів мережі, контролювати роботу, що виконується віддалено, і використовувати результати, отримані з мережі.

2. Мобільні пристрої можуть брати участь в мережі як постачальники ресурсів, а не лише як одержувачі послуг.

Важливо зауважити, що виникає багато проблемних питань, якщо розглядати мобільні пристрої як частину обчислювальних GRID ресурсів чи інтерфейсів. Одним з найбільш важливих є нестабільне підключення мобільних пристроїв. Аналогічні ситуації можна знайти в системі P2P (Peer-to-Peer), де процесорний цикл кожної машини, склад і компоненти можуть бути поділені для того, щоб розширити їх можливості [3]. У системі P2P, користувачі можуть вільно приєднуватись та залишати мережу, дослідження показало, що у середньому вузол залишається на зв'язку тільки 28% часу [4].

Постановка задачі

Часті відключення вузла значно погіршують обчислення, що компенсується за рахунок дублікації робочих завдань. Проте, ця стратегія непродуктивно витрачає велику кількість ресурсів і не може бути застосована у мобільних мережах, хоча вона гарантує поліпшення продуктивності. Важливо зауважити, що ця стратегія успішно застосовується у системах P2P, які мають достатню кількість ресурсів, однак, велика кількість дублікацій не може бути виконана у мобільній мережі тому що така система має порівняно невелику кількість вузлів та високу чутливість до розподілу ресурсів.

Для розв'язання цього питання необхідно розробити алгоритм формування стійкої структури мобільної GRID-системи з урахування нестабільності підключень мобільних вузлів. У зв'язку з тим, що мережа складається з певного

числа автономних локальних ресурсів, алгоритм формування черг брокера можна розділити на дві категорії:

- глобального планування роботи;
- локального планування роботи.

Важливо зауважити, що існує велика кількість алгоритмів розподілення та планування навантаження брокерів: на експлуатаційному рівні; на рівні продуктивності; на рівні економії ресурсів; на рівні даних.

Після того як брокер планує навантаження менеджери локальних ресурсів *MQS* (*Mobile Queuing Server*) планують локальні графіки навантажень. Найчастіше локальні менеджери використовують принцип *FIFO*, але деякі менеджери можуть використовувати більш складні алгоритми. Тому необхідно розробити алгоритм, який дозволить виконувати локальне планування, оскільки планування поширюється тільки на локальні мобільні ресурси. Найбільшу увагу при цьому необхідно зосередити на чутливості енергії і слабкої безпеки, а також на ненадійності мобільного навколишнього середовища. Часті відключення мобільних вузлів значно впливають на ефективність роботи системи, тому планування роботи здійснюється виходячи з очікуваного часу використання на різних мобільних вузлах. За наявності доступних мобільних вузлів, робота (навантаження) вилучається з черги завдань *MQS*, які необхідно зробити і доставляються до мобільного вузла і вводяться дані. Після завершення робіт, що виконуються у мобільних вузлах, вихідні дані повертається до *MQS* і мобільний вузол виконує наступне завдання із черги. У такій системі завдання можуть бути заплановані незалежно один від одного, що не потребує спілкування мобільних вузлів під час виконання завдання. Однак, мобільний вузол повинен бути в змозі з'єднатись під час передачі введених та виведених даних. Якщо мобільний вузол не в змозі з'єднатись, *MQS* чекає доки з'єднання відбудеться. На рис. 1 показано перехідну діаграму мобільного вузла.

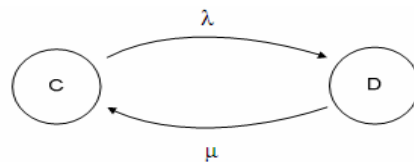


Рис. 1. Діаграма перехідного стану мобільного вузла

Пункти *C* і *D* позначають стан з'єднання (*C*) і стан відключення (*D*), відповідно. λ це стан роз'єднання та μ перепідключення мобільного вузла.

Важливо зауважити, що коли певна кількість завдань прибуває на *MQS*, або певна кількість робочих навантажень викликана однією роботою, *MQS* вибирає певні мобільні вузли, які виконують завдання. Коли мобільні пристрої беруть участь в обчисленні, з'єднання вузлів ненадійне, тобто вони можуть бути у стані відключення протягом тривалого часу, в той час як інші мобільні вузли працюють. Ненадійні мобільні вузли можуть знизити загальну продуктивність виконання роботи, тому що *MQS* повинна чекати підключення будь-якого відключеного вузла, для того, щоб завершити роботу. Також необхідно визначити, чи буде краще включити мобільний вузол в обчислення, чи ні.

Нехай $f_c(t)$ і $f_d(t)$ позначають очікуваний час, що необхідний для передачі даних (вхідних і вихідних) за одиницю часу t (передача даних потребує t одиниць часу у разі, якщо відключення немає), коли мобільний вузол спочатку знаходиться у стані з'єднання та стані роз'єднання, відповідно. Нехай P_C і P_D – ймовірність перебування стану зв'язку та стану його відключення, відповідно. Можемо отримати наступні рівняння перехідного стану, що показані на рис. 1:

$$P_C + P_D = 1 \quad \text{та} \quad -\lambda P_C + \mu P_D = 0.$$

Обчислимо за формулами:

$$P_C = \frac{\mu}{\mu + \lambda}, \quad P_D = \frac{\lambda}{\mu + \lambda}.$$

Тоді, $C(t, \varepsilon) = f_c(t + \varepsilon) - f_c(t)$ – час, що необхідний, щоб перетворити ε оди-

ниць часу для передачі даних, що почалася з часу t , коли мобільний вузол знаходився в стані зв'язку. Обчислимо $C(t, \varepsilon)$ у двох наступних випадках:

– відключення не відбувається при ε (ймовірність $e^{-\lambda\varepsilon}$): ε одиниці часу, що необхідні, для передачі даних.

– відключення відбувається при ε (ймовірність $1 - e^{-\lambda\varepsilon}$): $\frac{1}{\mu}$ одиниця часу, протягом якої очікується стан відключення.

Зауважимо, що ε одиниця часу, необхідна для передачі даних і $\frac{1}{\mu}$ – одиниця часу, яка в середньому потрібна на перепідключення. Отже, $C(t, \varepsilon)$ розрахуємо наступним чином:

$$\begin{aligned} C(t, \varepsilon) &= f_c(t + \varepsilon) - f_c(t) = \\ &= e^{-\lambda\varepsilon} \varepsilon + (1 - e^{-\lambda\varepsilon}) \left(\varepsilon + \frac{1}{\mu} \right). \end{aligned}$$

Для обчислення $f_c(t)$:

$$\begin{aligned} \frac{\partial f(t)}{\partial t} &= \lim_{\varepsilon \rightarrow 0} \frac{f(t + \varepsilon) - f(t)}{\varepsilon} = \\ &= \lim_{\varepsilon \rightarrow 0} \frac{e^{-\lambda\varepsilon} \varepsilon + (1 - e^{-\lambda\varepsilon}) \left(\varepsilon + \frac{1}{\mu} \right)}{\varepsilon} = \frac{\lambda}{\mu} + 1. \end{aligned}$$

Таким чином, ми отримаємо очікуваний час, що необхідний для передачі даних за t одиниць часу (передача даних триває t одиниць часу без відключення), коли мобільний вузол знаходиться у стані зв'язку, ми отримаємо наступне:

$$f_c(t) = \left(\frac{\lambda}{\mu} + 1 \right) t.$$

Очікуваний час, необхідний для передачі даних за t одиниць часу, коли мобільний вузол спочатку знаходиться в стані відключення обраховується наступним чином (Крім того, $\frac{1}{\mu}$ – одиниця часу, що потрібна в середньому для перепідключення.):

$$f_d(t) = \frac{1}{\mu} + f_c(t) = \frac{1}{\mu} + \left(\frac{\lambda}{\mu} + 1 \right) t.$$

Середній очікуваний час передачі даних, $f(t)$, є,

$$\begin{aligned} f(t) &= P_c \cdot f_c(t) + P_d \cdot f_d(t) = \frac{\mu}{\lambda + \mu} \cdot \left(\frac{\lambda}{\mu} + 1 \right) t + \\ &+ \frac{\lambda}{\lambda + \mu} \cdot \left\{ \frac{1}{\mu} + \left(\frac{\lambda}{\mu} + 1 \right) t \right\}. \end{aligned}$$

Якщо $f(t)$ функція від t , λ та μ ,

$$f(t, \lambda, \mu) = \left(\frac{\lambda}{\mu} + 1 \right) t + \frac{\lambda}{\lambda + \mu} \cdot \frac{1}{\mu}$$

Нехай t_{job} , t_{in} і t_{out} час, за який завдання обробляються мобільним вузлом, очікуваний час, необхідний для передачі вхідних даних до мобільного вузла, і очікуваного часу, необхідного для передачі вихідних даних з мобільного вузла, відповідно. Час виконання одного завдання мобільним вузлом I може бути представлений наступним чином:

$$g_i(t_{in}, t_{job}, t_{out}) = f_i(t_{in}) + t_{job} + f_i(t_{out}). \quad (1)$$

Алгоритм планування роботи для мобільної *GRID*-системи можна записати у наступному вигляді:

1. N := кількість мобільних вузлів в мобільному *GRID* об'єднанні ресурсів.

2. $total\ W$:= загальні витрати часу на виконання роботи.

3. Розрахунок часу, витраченого всіма мобільними вузлами, використовуючи рівняння.

4. Список мобільних вузлів у зростаючому порядку, залежно від часу виконання.

5. *FOR* ($i=N$; $i>1$; $i--$)

$$if \left[- (g_i - g_{i-1}) + \left(\frac{W_{total}}{i-1} - \frac{W_{total}}{i} \right) > 0 \right] Break;$$

6. Розподіляє роботу між вузлами i^{th} в списку.

Зауважимо, що у кроках 3 і 4, вузли перераховані в порядку зростання часу

виконання згідно рівняння (1). За крок 5, $g_i - g_{i-1}$ показує очікувані витрати додаткового часу на очікування, що спричинений включенням i^{th} вузла, потім $(i-1)^{th}$ вузол надає результат, а $\frac{W_{total}}{i-1} - \frac{W_{total}}{i}$ означає, що очікуване скорочення часу виконання відбулося через включення i^{th} вузла в обчислення. Таким чином, i^{th} вузол буде включений, якщо вдасться скоротити загальний час роботи в порівняно з включенням $(i-1)^{th}$ вузла.

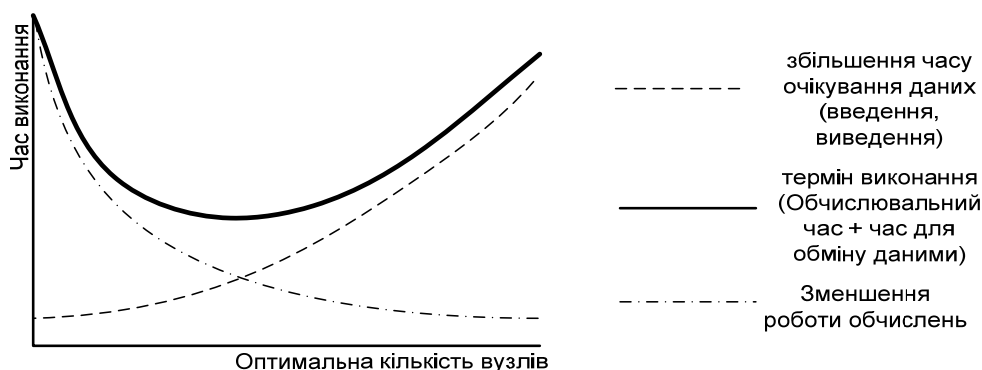


Рис. 2. Вигоди та втрати від збільшення кількості мобільних вузлів, що приймають участь в обчисленні

Висновки

У роботі запропоновано алгоритм формування стійкої структури мобільної GRID-системи з забезпеченням заданих параметрів якості обслуговування. Крім того враховано нестабільне підключення мобільних вузлів, що дало можливість підвищити продуктивність системи за рахунок подолання ненадійного підключення мобільного середовища, шляхом ретельного планування. Проте все ще існують складні проблеми, такі, як обмежена енергія, неоднорідність пристрою, безпека і так далі.

Список літератури

1. Foster I., Kesselman C., Tuecke S. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. – Supercomputer Applications, 2001. – Vol. 15(3). – P. 45–54.
2. Phan T., Huang L., Dulan C. Challenge: Integrating Mobile Wireless Devices

Якщо ми включимо більше вузлів в GRID обчислення, то кожен вузол виконує меншу кількість роботи, таким чином, час на обробку завдань мобільним вузлом зменшується (рис. 2). Однак, якщо ми включимо більшу кількість вузлів, найгірший вузол може бути в стані відключення довго, і це призведе до збільшення витрат часу на передачу вхідних і вихідних даних. Таким чином, необхідно знайти оптимальне число вузлів, що беруть участь, шукаючи компроміс між дефіцитом і ефективністю.

Into the Computational Grid. – MobiCom'02, 2002. – Atlanta, GA. – P. 512–518.

3. Ledlie J., Shneidman J., Seltzer M., Huth J. Scooped, Again. – IPTPS'03: Second International Workshop on Peer-to-Peer Systems, 2003. – Berkeley, CA, USA. – P. 310–372.

4. Wilcox-O'Hearn B. Experiences Deploying a Large-Scale Emergent Network. – IPTPS'02, 2002. – Cambridge, MA, USA. – P. 182–195.

Подано до редакції 15.03.10